

A mechanistic study of the spontaneous hydrolysis of glycyserine as the simplest model for protein self-cleavage

Tzvetan T. Mihaylov,^{*,[a]} Tatjana N. Parac-Vogt^[a] and Kristine Pierloot^[a]

Abstract

A common feature of several classes of intrinsically reactive proteins with diverse biological functions is that they undergo self-catalyzed reactions initiated by an N→O or N→S acyl shift of a peptide bond adjacent to a serine, threonine or cysteine residue. In this study we examine the N→O acyl shift initiated peptide bond hydrolysis at the serine residue on a model compound, glycyserine (GlySer), by means of DFT and *ab initio* methods. In the most favorable rate-determining transition state the serine –COO[−] group acts as a general base to accept a proton from the attacking –OH function, resulting in oxyoxazolidine ring closure. The calculated activation energy (29.4 kcal/mol) is in excellent agreement with the experimental value, 29.4 kcal/mol, determined by ¹H NMR measurements. A reaction mechanism for the entire process of GlySer dipeptide hydrolysis is also proposed. In the case of proteins we found that, when no other groups that may act as a general base are available, the N→O acyl shift mechanism might instead involve a water-assisted proton transfer from the attacking serine –OH group to the amide oxygen. However, the calculated energy barrier for this process is relatively high (33.6 kcal/mol), thus indicating that in absence of catalytic factors the peptide bond adjacent to serine is no longer a weak point in the protein backbone. An analogous rearrangement involving the amide N-protonated form rather than the principle zwitterion form of GlySer was also considered, as a model for the previously proposed mechanism of SEA domain autoproteolysis. The calculated activation energy, 14.3 kcal/mol, is significantly lower than the experimental value reported for SEA (~21 kcal/mol), but is still in better agreement as compared to earlier theoretical attempts.

Keywords: glycyserine, peptide bond hydrolysis, reaction mechanisms, density functional calculations, *ab initio* calculations, protein self-cleavage

Address:

[a] Dr. T. T. Mihaylov, Prof. Dr. T. N. Parac-Vogt, Prof. Dr. K. Pierloot
KU Leuven, Department of Chemistry
Celestijnenlaan 200F, B-3001 Leuven, Belgium.
Fax: +32-16-327992
E-mail: tzmihay@svr.igic.bas.bg

Introduction

Uncatalyzed hydrolysis of peptide bonds is an extremely slow reaction that proceeds with a half-life of about 600 years at 25 °C and neutral pH.^[1] In contrast to the evolutionary selection for chemically stable peptides, there are several classes of intrinsically reactive proteins that cleave the internal peptide bond as part of their cellular function, without the need of an auxiliary enzyme or cofactor.^[2] Typical examples of self-cleaving proteins are the intein domain in protein splicing,^[3] the pyruvoyl-dependent enzymes^[4], the *hedgehog* domain,^[5] the N-terminal nucleophile (Ntn) hydrolases,^[6] the nucleoporins^[7], and the SEA (sea urchin sperm protein, enterokinase, and agrin) domain.^[8] Although the biological functions of these proteins are diverse, there are striking similarities between them when it comes to the mechanism of autoprolysis. In all of them, the self-catalyzed reactions are initiated by an N→O or N→S acyl shift (also called acyl transfer, migration, or rearrangement) converting the scissile peptide bond to a (thio)ester.^[2] The mechanism of this process involves the intramolecular nucleophilic attack by an –OH or –SH group of a serine (Ser), threonine (Thr) or cysteine (Cys) residue at the adjacent amide carbon of the protein backbone, generating an oxyxazolidine or an oxythiazolidine anion intermediate, respectively. This transient species collapses to a linear (thio)ester which is further resolved by a subsequent reaction that differs between autoprolytic systems (Figure 1).^[9] At physiological conditions the equilibrium of an N→X (X = O, S) acyl shift lies in favor of the amide bond, thus limiting the peptide bond cleavages.^[10] However, the protein environment around the scissile bond may significantly influence this process. Several factors accelerating the formation of a (thio)ester have been identified in self-cleaving proteins: deprotonation of the nucleophilic –XH group,^[3, 11] stabilization of the transient tetrahedral intermediate via an oxyanion hole;^[11a, 12] protonation of the leaving amino group,^[11b, 11c] and ground-state destabilization of the scissile peptide bond through conformational strain.^[8, 13]

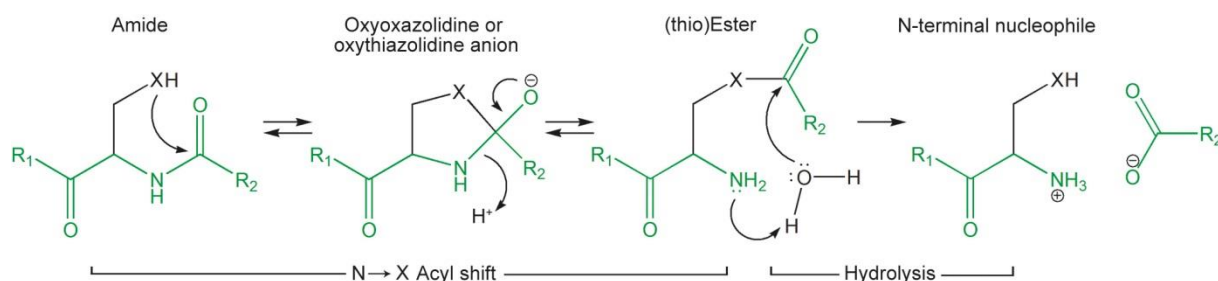


Figure 1. General scheme of peptide bond hydrolysis initiated by an N→X acyl shift (X = O, S).

To understand the extent to which catalysis is required for the self-cleavage of proteins Lewis and Wolfenden examined the hydrolysis of two model compounds, N-acetylserine and N-acetylcysteine, where the aforementioned rate enhancing factors are absent.^[14] They found that, as compared to the rate of hydrolysis of N-acetylalanine in neutral solution at 25 °C, the –SH group of N-acetylcysteine and the –OH group of N-acetylserine enhance the rate of their hydrolysis by a factor of 70 and 12 respectively.^[14] In our previous studies, the degree of peptide bond hydrolysis of a series of dipeptides was determined at 60 °C and neutral pH.^[15] Among all peptides examined, those with a Ser or Thr residue at the C-terminus were most rapidly hydrolyzed (Cys-containing dipeptides were not examined). These results clearly

demonstrate that the peptide bond upstream to Ser (Thr) and Cys residues are inherently susceptible to decomposition in physiological conditions.

For proteins, nucleophilic attack by an -XH group on the amide carbon is generally accepted to initiate autoproteolysis. However, the molecular hydrolysis mechanism of small model compounds, such as N-acetylated amino acids or dipeptides, is still controversial. For N-acetylated Ser and Cys it has been suggested that the rate-enhancing effect of the -XH group arises from its ability to act as an intramolecular general acid-base catalyst for hydrolysis (the deprotonated -XH group assisting the addition of a water molecule on the amide carbon), rather than as a nucleophilic catalyst.^[14] Hence, it was assumed that the specific protein environment of the scissile peptide bond should play a decisive role, redirecting the actions of the side chains to nucleophilic attack.^[14] This suggestion seems reasonable in the case of acetylcysteine ($pK_a(\text{SH}) = 9.52$ ^[16]), where the concentration of deprotonated (-S⁻) form at neutral pH could be sufficient to tip the balance from a nucleophilic to a general base mechanism. However, it is hardly acceptable in the case of acetylserine, since the alcoholic hydroxy group of serine is such a weak acid ($pK_a(\text{OH}) = 14.4$ ^[17]) that it does not undergo dissociation at neutral pH. A general base mechanism also cannot explain the fact that peptide bond hydrolysis is much faster for GlySer than for SerGly.^[15] A reasonable explanation of this phenomenon can be provided, however, by a nucleophilic mechanism. Indeed, intramolecular attack of the Ser hydroxy group on the amide carbon is possible in GlySer, giving rise to a five-membered ring transition state. Such an attack is, however, impossible in SerGly, because it would in this case lead to an unfavorable four-membered ring transition state.

An N→O acyl shift mechanism (Figure 1) has also been suggested for the hydrolysis of dipeptides with a Ser residue at the C-terminus, which is promoted by metal cations^[18] (Zn²⁺, Ni²⁺, and Cd²⁺ salts) as well as by oxometalate anions^[15] (H₂VO₄⁻, MoO₄²⁻) at neutral pH. In both cases, the hydrolytic effect of the metal has been attributed to its ability to coordinate the peptide carbonyl O-atom, thus enhancing the susceptibility of the amide C-atom toward internal nucleophilic attack by the -OH group of the Ser residue. DFT calculations support this hypothesis,^[15b] but a thorough theoretical investigation of the hydrolysis mechanism has so far not been performed.

Several theoretical studies have already revealed important insights into the self-catalyzed N→X acyl shift mechanism in proteins. They are mainly devoted to protein splicing and SEA domain autoproteolysis, where the process is accelerated either through deprotonation of the nucleophilic -XH group/protonation of the leaving amino group^[11b] or through conformational strain of the peptide bond.^[8b, 19] In contrast, the molecular mechanism of the uncatalyzed peptide bond cleavage (via an N→X acyl shift) of small exemplary model compounds, such as Ser(Thr)- or Cys-containing dipeptides, has so far not yet been investigated.

To examine the inherent hydrolytic ability of serine-containing peptides, we have performed a detailed mechanistic study of the spontaneous hydrolysis of GlySer in neutral aqueous solution by means of DFT and *ab initio* methods. The role of the solvent in this process is elucidated by considering the “most important” water molecules explicitly, and adopting a cluster-continuum model. We hope that this work will also serve as a valuable basis for a better understanding of more complicated cases, such as the metal and oxoanion promoted hydrolysis of serine-containing dipeptides.

Computational methods

The reaction mechanism of GlySer hydrolysis in neutral aqueous solution was modeled by explicit inclusion of water molecules in the vicinity of the peptide and implicit treatment of

the rest of the solvent with a continuum solvation model (i.e. a mixed cluster-continuum treatment). Such an approach has been shown to be appropriate for studying organic and organometallic reactions in solution. Many examples were collected in a recently published review of Sunoj and Anand.^[20] Experimental evidence revealing the important role of the water media in reactions analogous to the N→O acyl shift in GlySer were previously reported.^[21]

According to the cluster-continuum concept, the number of explicit solvent molecules (n) in the transition state (TS) has to be varied in order to minimize the activation energy of a chemical reaction.^[22] Following this strategy, extensive molecular modeling of transition states, with n ranging from 0 to 4 was carried out. Because of the increasing conformational freedom of the system when increasing n , this computationally expensive procedure was applied only for the rate-limiting TSs of the corresponding reaction channels and the rest of the reaction processes were modeled with one explicit water molecule. However, all of the presented and compared TS structures in this work are the lowest energy structures found for the given TS mechanism and n . The arrangement of the water molecules in every lowest energy TS. n H₂O cluster is the result of a systematic search and optimization of many different models.

The molecular geometries of all species modeled here were optimized in solution without constraints using the M06 functional^[23] and making use of 6-311+G(2df,2p) basis sets, conjugated with an SMD^[24] continuum solvation model. The same basis sets were also employed in single point energy calculations with the post-HF method MP2 and its spin-component-scaled version SCS-MP2,^[25] as well as with the double-hybrid functionals mPW2PLYP-D^[26] and B2PLYP-D^[27], incorporating an empirical dispersion correction.^[28] To confirm the character of all first-order saddle points and local minima on the potential energy surfaces (PESs) vibrational frequency calculations were carried out. Intrinsic reaction coordinate (IRC) calculations were performed to verify the expected connections between the first-order saddle points and local minima on the PES.

To calculate the Gibbs free energy of each species in solution the zero-point vibrational energy, thermal correction and entropy term were obtained from the aqueous phase M06 frequency calculations by using the standard rigid rotor-harmonic oscillator approximation without scaling. In most previous theoretical studies, frequency calculations on gas phase optimized structures were adopted for the evaluation of free energies.^[29] However, the zwitterion form, which is the major form of the amino acids in neutral aqueous solution, is not stable in gas phase and most of the transition states and intermediate structures studied here cannot be located via gas phase calculations (or at least would require a large number of explicit solvent molecules in gas phase) because of the loss of bulk solvent effects. Direct calculations of the Gibbs energy in solution were also performed in other recent studies, and were shown to be a practical approach which usually yields reasonably accurate results.^[30] The correctness of this approach, however, has been questioned by Ho et al.^[31] As a response, Ribeiro et al. have shown in their study that using partition functions computed for molecules optimized in solution is a correct and useful approach for averaging over solute degrees of freedom when computing free energies of solutes in solution, and moreover, it is recommended for cases where liquid and gas-phase solute structures differ appreciably or when stationary points present in liquid solution do not exist in gas phase.^[32]

The free energy of an arbitrary compound X in aqueous phase may be expressed as a sum of its gas-phase free energy $G^0(X)$ and the free energy of hydration $\Delta G_{\text{sol}}^*(X)$.

$$G_{\text{aq}}(X) = G^0(X) + \Delta G_{\text{solv}}^*(X) \quad (1)$$

Since the $G^0(X)$ values are usually calculated with respect to a standard state of 1 atm (denoted by the superscript “⁰”), as is the case here, whereas the SMD hydration free energies refer to a standard state of 1 M^[24] (denoted by the superscript “*”), a correction term $\Delta G^{0 \rightarrow *} = RT \ln(24.46)$ should be added to the calculated free energy to convert to a standard state of 1 M.

$$G_{\text{aq}}^*(X) = G_{\text{aq}}(X) + \Delta G^{0 \rightarrow *} \quad (2)$$

At 298.15 K, $\Delta G^{0 \rightarrow *}$ is 1.89 kcal/mol. The standard state concentration of H₂O in the liquid phase is 55.34 M. In order to account for this an additional concentration correction term $RT \ln([H_2O])$, corresponding to the free energy change of 1 mol of H₂O gas from the 55.34 M liquid state to the 1 M standard state in solution, should be included in the calculation of the standard free energy of H₂O in the aqueous phase. At 298.15 K, $RT \ln([H_2O]) = 2.38$ kcal/mol. For more details the reader is referred to Ref. [33] and [34]. Thus, at a standard state of 1 M, taking the energy of the separated reactants (GlySer and n .H₂O) as a reference the relative aqueous-phase Gibbs energy of a stationary point X (involved in the reaction of GlySer hydrolysis) microsolvated by n water molecules, $\Delta G_{\text{aq}}^*(X(H_2O)_n)$, could be expressed as

$$\Delta G_{\text{aq}}^*(X(H_2O)_n) = G_{\text{aq}}(X(H_2O)_n) - G_{\text{aq}}(\text{GlySer}) - n(G_{\text{aq}}(H_2O) + RT \ln([H_2O]) + \Delta G^{0 \rightarrow *}) \quad (3)$$

The reaction free energy of GlySer hydrolysis in neutral aqueous solution is calculated according to the equation

$$\Delta G_{\text{aq,react}}^* = G_{\text{aq}}(\text{Gly}) + G_{\text{aq}}(\text{Ser}) - G_{\text{aq}}(\text{GlySer}) - G_{\text{aq}}(H_2O) - RT \ln([H_2O]) \quad (4)$$

All calculations reported in this paper were performed with the program package Gaussian 09.^[35]

Experimental section

The spontaneous hydrolysis of 2.0 mM solution of GlySer at pD 7.4 and pD 5.4 was followed at 60, 70, 80 and 90 °C. The observed rate constants (k_{obs}) were calculated by integration of the ¹H NMR signal of Gly, one of the hydrolysis products, at different time increments and by plotting its concentration as a function of time. The ¹H NMR spectra were recorded with a Bruker Advance 400 spectrometer. Activation Gibbs function parameters (ΔG^\ddagger , ΔH^\ddagger and ΔS^\ddagger) were obtained from the Arrhenius plot.

Results and discussion

The first part of the discussion elaborates on the possible mechanisms of the first step of the N→O acyl rearrangement, which is the bottleneck of the overall spontaneous GlySer hydrolysis in neutral aqueous solution. On the basis of energetic comparisons the most favorable pathway for the rate-limiting step is identified. In the second part, a detailed molecular mechanism of the entire process of hydrolysis is proposed and discussed. In the third part, selected transition states that could be relevant for the self-cleavage of serine-

containing peptides are discussed, thus making a link between the model compound and the protein backbone.

Intramolecular nucleophilic attack by the serine -OH group

As expected and confirmed by the calculations, the nucleophilic attack initiating the N→O acyl rearrangement has the highest energy requirement, thus constituting the rate-limiting step of the overall GlySer hydrolysis. Therefore, our efforts to model different possible transition states and to minimize their free energies by gradually changing the number of explicit solvent molecules, were primarily focused on this part of the reaction.

The nucleophilic addition of the alcoholate oxygen on the amide carbon, leading to the formation of a transient oxyxazolidine ring, requires deprotonation of the serine -OH group. This process could take place in a stepwise manner or in a concerted manner. The most obvious atomic sites in the GlySer zwitterion (the predominant form of GlySer at neutral pH) that might serve as potential acceptors of the -OH proton are the carboxyl oxygens, the amide nitrogen and the amide oxygen. The corresponding reaction mechanisms are schematically presented in Figure 2, and denoted as Channel A, Channel B and Channel C respectively.

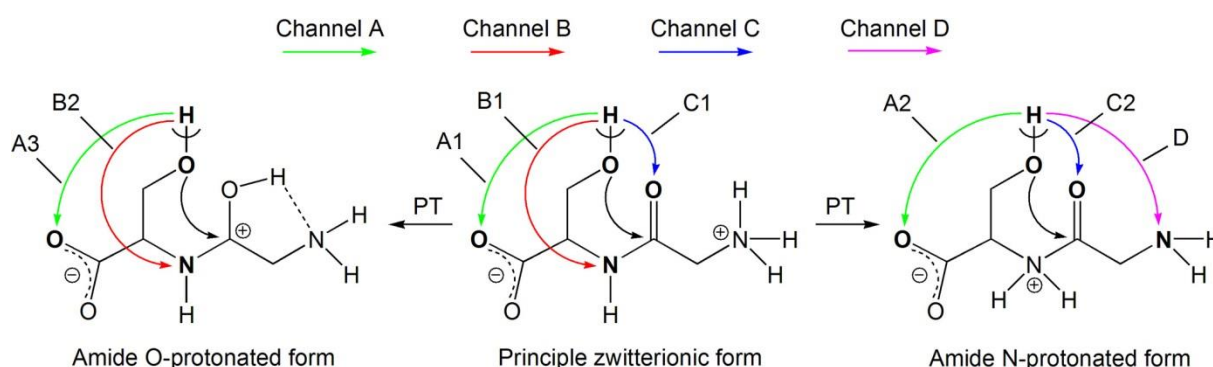


Figure 2. Reaction channels for intramolecular nucleophilic attack by the serine -OH group on the amide carbon. Three possible GlySer starting structures are considered: the principle zwitterionic form, the amide N-protonated form and the amide O-protonated form. The last two forms are obtained from the principle zwitterion one through a PT from the -NH_3^+ group to the amide nitrogen or to the amide oxygen respectively.

Besides the principle zwitterionic form, analogous reaction mechanisms were modeled by using two more GlySer isomers, the amide N-protonated form and the amide O-protonated form. In those cases we assume that an intramolecular proton transfer (PT) from the -NH_3^+ group to the amide nitrogen or to the amide oxygen could precede the actual nucleophilic attack by the Ser -OH group. In the former case the calculated energy barrier related to the PT is 26.9 kcal/mol. The resultant amide N-protonated form is a high energy local minimum structure with a relative free energy (24.2 kcal/mol) that is only 2.7 kcal/mol below the energy barrier for its formation. In the latter case, the amide O-protonated form was located as minimum on the PES, but it is not a minimum on the free energy surface. Thus, it has a relative free energy of 10.5 kcal/mol, while an energy requirement of 9.8 kcal/mol is calculated for the PT. Therefore, the amide O-protonated form should not be considered as a

stable starting structure and the PT related to its formation should be viewed as an integral part of the TS of a -OH addition (see also Figure S1).

Ultimately, as shown in Figure 2, several variants of channel A (A1, A2 and A3), channel B (B1 and B2) and channel C (C1 and C2) were examined. As an alternative to the proton accepting sides involved in these three channels the amide N-protonated form of GlySer has a free amino group which also could uptake a proton from the attacking OH nucleophile. This mechanism is referred to as channel D. Our calculations revealed that in all studied mechanisms the serine -OH deprotonation and the ring closure proceed in a concerted manner. Depending on the starting GlySer isomer and the reaction channel the TS of the addition reaction collapses in an oxyoxazolidine (A1), hydroxyoxazolidine (A3, C1), N-protonated oxyoxazolidine (A2, B1, D) or N-protonated hydroxyoxazolidine (B2, C2) ring intermediate.

Based on the arguments presented in the introductory part the general base mechanism of hydrolysis has been ruled out. Nevertheless, we have tried to model the transfer of a proton from the serine -OH group to the bulk solvent, thus barring the -O^- as a precondition of its general base activity. Even in models where the H_3O^+ ion was stabilized by three water molecules neither a transition state nor a stable intermediate, connected to such a TS, could be located on the PES. In the course of the optimization, the proton in all cases spontaneously moved back to the alcoholate oxygen of serine.

Next, the four channels for nucleophilic attack, A, B, C, and D, will be discussed in detail.

Channel A. Within this reaction mode three different possibilities, A1, A2 and A3, were examined. The corresponding lowest energy transition states, with the free energy of activation as a function of the number of water molecules (n) are depicted in Figure 3.

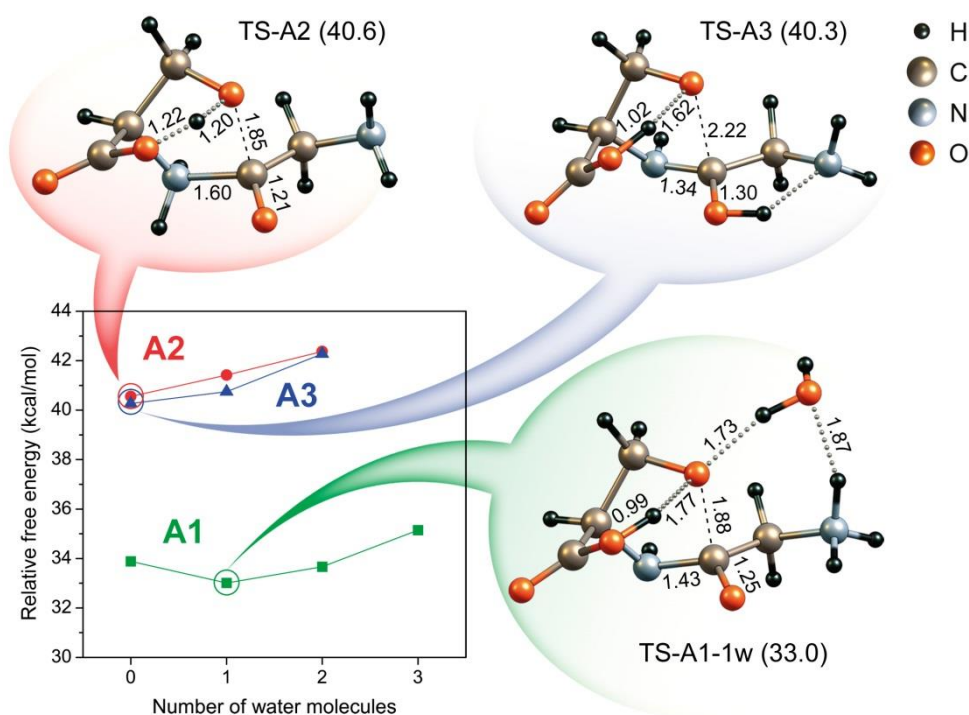


Figure 3. SMD-M06/6-311+G(2df,2p) optimized geometries and free energies of activation for the lowest energy transition states of intramolecular nucleophilic attack via A1, A2 and A3 reaction channels. Energies (in brackets) are in kcal/mol, interatomic distances are in Å.

The activation free energy for oxyoxazolidine ring closure through channel A1 reaches its minimum when the transition state is stabilized by one water molecule (TS-A1-1w). In the TS-A1-1w structure the serine –OH proton is already fully transferred to the carboxyl oxygen (O-H distance = 0.99 Å) and the ring is yet to close (O-C distance = 1.88 Å). The process is assisted by hydrogen bonding of one water molecule to both the alcoholate oxygen and a -NH₃⁺ hydrogen. The calculated activation free energy of channel A1 is 33.0 kcal/mol with the M06 functional. In contrast, channels A2 and A3 (transition states TS-A2 and TS-A3, respectively) do not require specific assistance of solvent molecules. In the TS-A2 structure the H⁺ is found halfway between the two oxygens and the nascent –O⁻ attacks the amide carbon at a distance of 1.85 Å. This is not the case, however, for TS-A3 (channel A3), where the –OH proton is almost completely transferred to the adoptive oxygen (O-H distance = 1.02 Å). As mentioned above, the amide O-protonated isomer is not a minimum energy structure (on the basis of the free energy), such that its formation should be considered as an integral part of the transition state. Thus, in the end two proton transfers accompanied the ring closure via the A3 mechanism. Although the starting structures and the attacking distances are different, the free energies of activation of the A2 and A3 pathways are rather close, 40.6 and 40.3 kcal/mol respectively.

Optimized geometries of the higher energy transition states, not depicted in Figure 3, may be found in Figure S1 (in the Supporting information).

Channel B. Within this reaction channel two types of transition states, TS-B1 and TS-B2, were modeled. Relevant transition state structures, energies and interatomic distances are given in Figure 4. Additional TS structures may be found in Figure S3. The intramolecular nucleophilic attack via the B1 mechanism requires a bridge consisting of two water molecules, (TS-B1-2w) mediating the PT from the serine -OH to the amide nitrogen. In TS-B1-2w the new O-C bond is almost fully formed (O-C distance = 1.50 Å) and the -OH proton is already transferred to a bridging water to form a hydronium ion. The calculated activation energy for channel B1 is 41.9 kcal/mol.

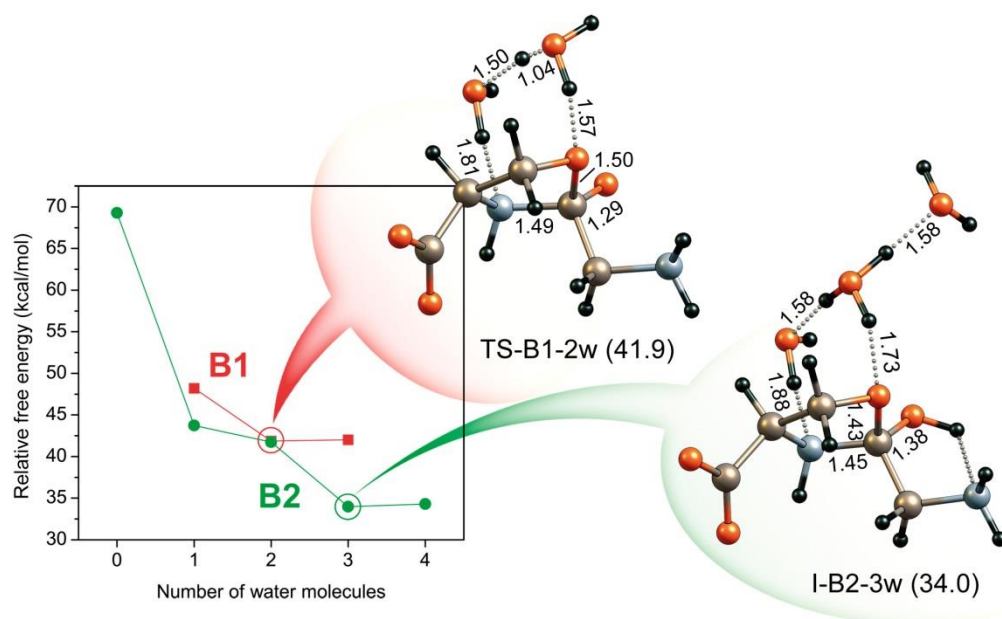


Figure 4. SMD-M06/6-311+G(2df,2p) optimized geometries and free energies of activation for the lowest energy transition states of intramolecular nucleophilic attack via B1 and B2 reaction channels. Energies (in brackets) are in kcal/mol, interatomic distances are in Å. For B1 we did not succeed in obtaining an appropriate transition state where a direct, rather than a water mediated, proton transfer takes place.

For the B2 channel, we found a minimum in the activation energy when three water molecules participate in the PT process. Structure I-B2-3w in Figure 4 was found as local minimum on the PES connecting two TSs corresponding (i) to a PT from the serine –OH group to a neighboring water molecule (involved in a bridge) with a formation of H_3O^+ and (ii) to a PT from the resultant H_3O^+ to the amide nitrogen. However, according to the free energy calculations this intermediate lies higher than both the preceding and the sequel TS. Thus, the B2 channel must be considered as a complex concerted process starting with a PT from the $-\text{NH}_3^+$ group to the $\text{C}=\text{O}$ oxygen, followed by a nucleophilic attack of the serine –OH on the amide carbon with simultaneous formation of a hydronium ion, and finally the decomposition of this ion with a water mediated PT to the amide nitrogen. During this process, a third water molecule stabilizes the hydronium ion through the formation of a strong hydrogen bond ($\text{H}_{\text{hydr}} \cdots \text{O}_{\text{w}}$ distance = 1.58 Å). The activation free energy for the B2 channel is found to be lower by 7.9 kcal/mol than for the B1 channel.

Channel C. Within this reaction mode two pathways, C1 and C2, were considered. The lowest energy TS structures are depicted in Figure 5. The rest of the TSs modeled may be found in Figure S4.

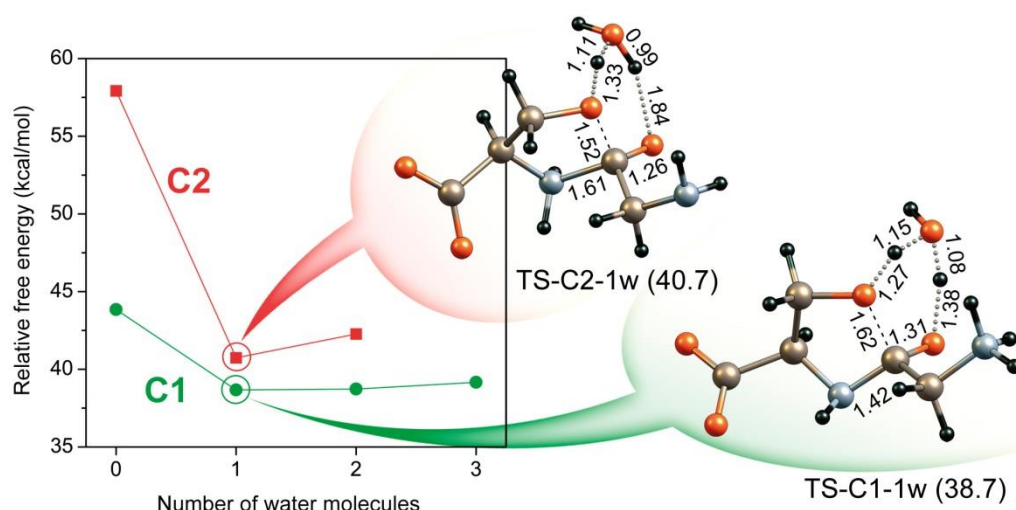


Figure 5. SMD-M06/6-311+G(2df,2p) optimized geometries and free energies of activation for the lowest energy transition states of intramolecular nucleophilic attack via C1 and C2 reaction channels. Energies (in brackets) are in kcal/mol, interatomic distances are in Å.

The energy barrier for the rate-limiting step of GlySer hydrolysis via channel C1 minimizes when the PT from the serine residue to the amide oxygen is mediated by one water molecule. Experimental evidence for participation of water molecules forming catalytic proton bridges in mechanism analogous to C1 were previously reported for a system that mimics serine protease enzymes.^[21] According to our calculations, such a bridge favors the nucleophilic addition by 6.3 kcal/mol as compared to the non-mediated PT, while inclusion of a second water molecule in the bridge raises the activation energy by 1.4 kcal/mol. In the transition state TS-C1-1w, there are two protons in flight and the alcoholate oxygen attacks the amide carbon at a distance of 1.62 Å. TS-C2-1w is very similar to TS-C1-1w, but here the computed attacking distance is shorter, 1.52 Å, probably due to the loss of planarity of the peptide bond in the amide N-protonated isomer. Also, the energy lowering of the water mediated PT with respect to the non-mediated case is now 17.2 kcal/mol. The calculated free energy of TS-C1-1w is lower by 2.0 kcal/mol than for TS-C2-1w.

Channel D. The free energy of activation for channel D minimizes when the solute is hydrogen bonded by three explicit solvent molecules, as shown in Figure 6. The complete set of TS structures modeled is provided in Figure S5.

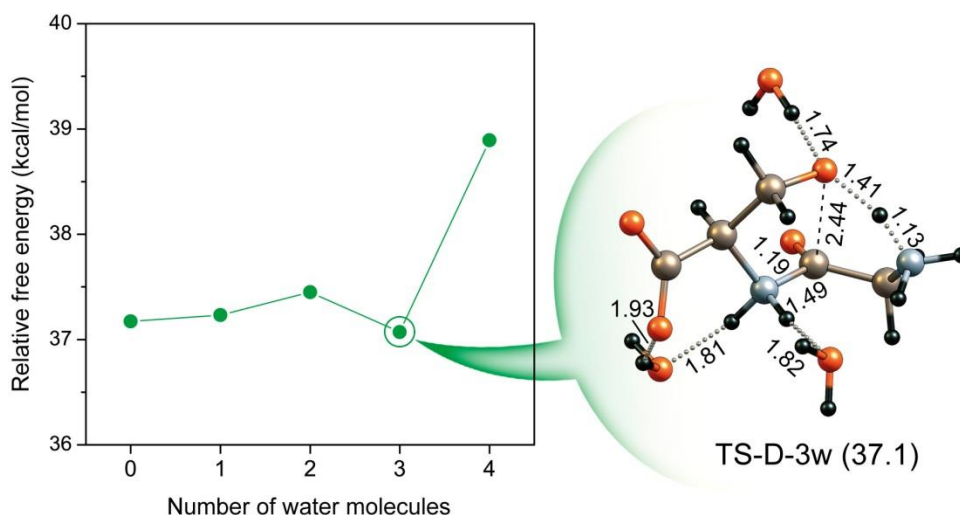


Figure 6. SMD-M06/6-311+G(2df,2p) optimized geometry of the lowest energy transition state and free energy of activation as a function of the number of water molecules for the intramolecular nucleophilic attack via channel D. Energy (in brackets) is in kcal/mol, interatomic distances are in Å.

According to our calculations, direct PT from the serine residue to the amino nitrogen is favored by 4.7 kcal/mol with respect to a water mediated PT. In comparison with all other transition states described above, the lowest energy transition state of channel D, TS-D-3w, reveals the largest distance between the nascent O^- and the amide carbon (2.44 Å). Its activation free energy, 37.1 kcal/mol, is the lowest when compared to the other TSs based on an amide N-protonated isomer (TS-A2: 40.9 kcal/mol, and TS-C2-1w: 40.7 kcal/mol).

Activation energy of the GlySer hydrolysis. The activation parameters of the spontaneous hydrolysis of GlySer were first experimentally determined (this work) at physiological pH (pD 7.4). However, at this pH the concentration of the anionic form of GlySer is significant, 11.5 % ($\text{pK}_a(\text{NH}_2) = 8.38$).^[16] The presence of this anion could influence the hydrolysis reaction through a competitive (to the zwitterionic) mechanism, which in turn may be reflected in the observed activation parameters. To check whether the anion plays an important role in the hydrolysis mechanism at neutral pH the experimental activation parameters were also determined at pD 5.4 (the isoelectric point of GlySer is 5.68), where only the zwitterion is presented in the reaction solution. The results are shown in Table 1. Although the activation Gibbs functions estimated at pD 7.4 and at pD 5.4 are virtually the same the activation entropies and enthalpies exhibit significant differences. These results strongly suggest that the mechanism of the rate-limiting step of hydrolysis at neutral pD and at pD 5.4 might be thoroughly different. This issue will be discussed in detail in a separate publication. For the current study we have chosen to compare our calculated results for the hydrolytic activity of the GlySer zwitterion with the experimental activation parameters obtained at pD 5.4.

Table 1. Experimental activation parameters (energies in kcal/mol, entropy in cal/mol K) of the hydrolysis of 2.0 mM GlySer at pD 7.4 and pD 5.4.

pD	E_a	ΔH^\ddagger	ΔS^\ddagger	$\Delta G^\ddagger_{298\text{ K}}$
7.4	21.4	20.7	-27.0	28.7
5.4	26.3	25.6	-12.7	29.4

From a comparison of the calculated activation free energies for the various reaction channels A–D presented above, we may conclude that the rate-limiting step of GlySer zwitterion hydrolysis most probably proceeds through channel A1. The energy of TS-A1-1w, 33.0 kcal/mol, is the lowest among all TSs considered. An alternative pathway might be channel B2 with an activation energy (34.0 kcal/mol) just 1 kcal/mol above channel A1. All other reaction channels revealed much higher transition state energies, in the range of 37.1 – 41.9 kcal/mol. Although the calculated activation free energy of 33.0 kcal/mol (M06/6-311+G(2df,2p)) is in reasonable agreement with the experimental value (29.4 kcal/mol), the accuracy of the calculated reaction energetics might be further improved by using more appropriate computational methods. To investigate this, we have performed single point energy calculations on the A1 reaction channel with the more accurate double-hybrid density functionals (B2PLYP-D and mPW2PLYP-D) and with post-HF methods (MP2 and SCS-MP2). The results are summarized in Table 2, and compared to available experimental data for the reaction free energy of peptide bond hydrolysis in dipeptides^[36] and to the activation free energy from Table 1 (pD 5.4).

Table 2. Reaction and activation free energies (kcal/mol), calculated with different theoretical methods.

Method	$\Delta G_{aq,react}^*$	$\Delta G_{aq}^{*,\ddagger}$	Abs. error ^b
M06	-1.8	33.0	3.6
B2PLYP-D	-1.5	31.4	2.0
mPW2PLYP-D	-1.5	31.6	2.2
MP2	0.2	29.4	0.0
SCS-MP2	-0.6	32.8	3.4
Experiment	-3.6 ^a	29.4	

^a Since the reaction free energy of GlySer hydrolysis is not available, for sake of comparison, here we use the experimental reaction free energy of GlyGly hydrolysis, taken from Ref [36].

^b Absolute error of the calculated activation free energy of TS-A1-1w.

Both the double-hybrid and MP2 methods produce activation free energies that are in closer agreement with experiment than the M06 results, with absolute errors in the range of 0.0 – 3.4 kcal/mol. Notably, the MP2 calculated barrier for the rate-determining step (29.4 kcal/mol) is in excellent agreement with the observed activation Gibbs function (29.4 kcal/mol, at 298 K). This is not the case, however, for the reaction free energy, which is most accurately reproduced with the M06 functional (absolute error of 1.8 kcal/mol), while the largest deviation from experiment is obtained with the MP2 method (absolute error of 3.8 kcal/mol). However, given that in a previous study of a series of gas-phase reactions involving main elements, a root mean square error of 4.6 kcal/mol was reported for the MP2 reaction energetics^[25], the error obtained with this method in this work may still be considered reasonable. It should be taken also into account that the experimental value corresponds to the

reaction free energy of GlyGly as an approximation of $\Delta G_{aq,react}^*$ of GlySer (which is not known). The double-hybrid methods B2PLYP-D and mPW2PLYP-D do reveal a more stable behavior, showing a systematic error of about 2 kcal/mol for both the activation and the reaction (free) energies.

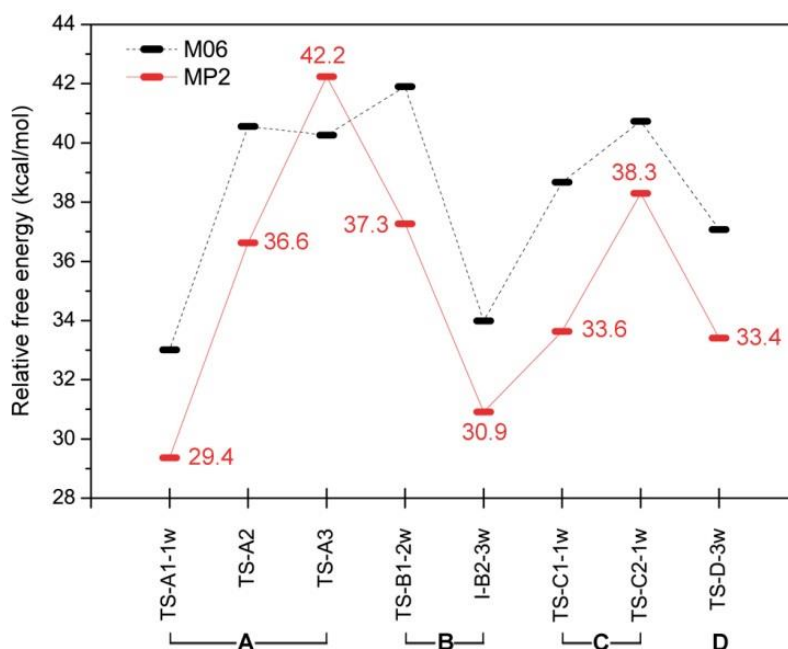


Figure 7. Comparison of the transition state free energies of channels A, B, C, and D, calculated with the M06 functional and the MP2 method.

As the primary goal of this work is to obtain reliable information concerning the hydrolysis mechanism and the associated energy barriers we have chosen the MP2 method (i.e. the best in reproducing the experimental activation energy) for a further refinement of the energy profile of GlySer hydrolysis with respect to M06. Hereto, the free energies of all TSs, representing the reaction channels A-D, were recomputed with the MP2 method. The results are graphically presented in Figure 7. With exception of TS-A3, all MP2 TS energies are shifted to lower values by 2.4 (TS-C2-1w) to 5.0 (TS-C1-1w) kcal/mol, as compared to their M06 values. The MP2 free energy difference between the most probable candidates, TS-A1-1w and I-B2-1w, is 1.5 kcal/mol in favor of channel A1 (as compared to 1.0 kcal/mol with M06). Unless otherwise noted, our discussion below will be based on the MP2 free energies.

Proposed reaction mechanism for the entire process of GlySer hydrolysis

The optimized structures of all species involved in the spontaneous hydrolysis of GlySer are shown in Figure 8 and the corresponding energy profile is depicted in Figure 9.

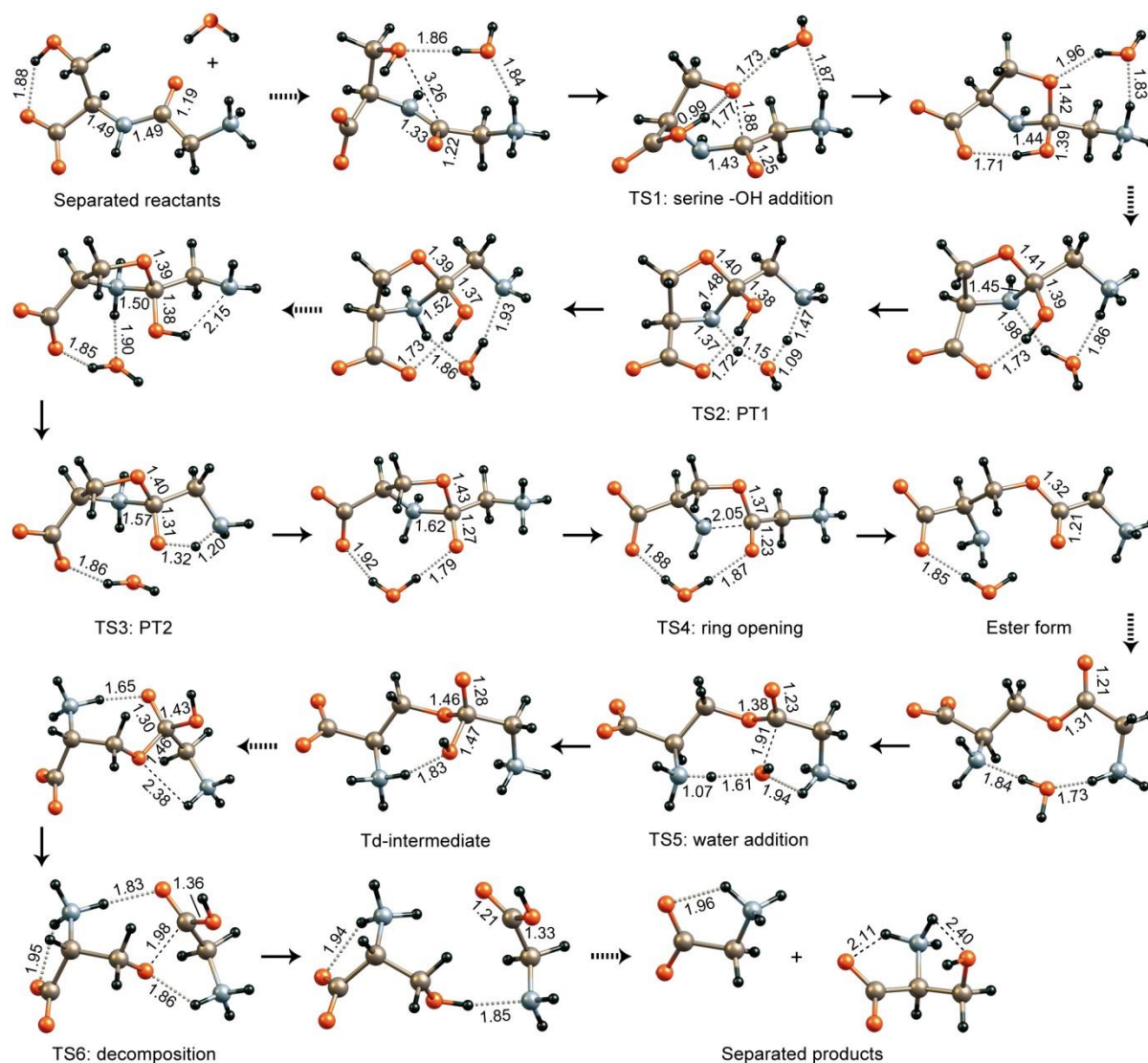


Figure 8. Proposed reaction mechanism for the spontaneous hydrolysis of GlySer in neutral aqueous solution. Relevant interatomic distances are given in Å. The structures obtained by following intrinsic reaction coordinates are connected with a solid arrow. Conformational interconversions^[37] are indicated with a dashed arrow.

The first barrier to overcome in channel A1, TS-A1-1w (already described above) is 29.4 kcal/mol. Following the path downhill from this state, a new O-C bond (ring closure) is formed, with a simultaneous proton transfer from the carboxyl group to the nascent oxyoxazolydine oxygen. The resultant hydroxyoxazolydine intermediate lies higher than the initial reactants by 12.2 kcal/mol. Decomposition of this intermediate with formation of a new -NH₂ terminus requires protonation of the ring nitrogen (see Figure 1). However, our (M06) calculations indicated that a water mediated PT directly from the hydroxyoxazolydine oxygen to the ring nitrogen would cost 34.5 kcal/mol. Therefore, an alternative pathway was modeled, involving two lower energy PTs (designated as PT1 and PT2). PT1 involves an H⁺ transfer from the terminal -NH₃⁺ group to the ring nitrogen via a bridging water molecule, while the NH₃⁺ form is recovered in the subsequent PT2 by means of a direct H⁺ transfer from the N-protonated hydroxyoxazolydine oxygen to the amino nitrogen. The transition states of both PT1 and PT2 have a relative energy of 22.7 kcal/mol with respect to the initial reactants. The

resulting intermediate (at 17.9 kcal/mol, with a net transfer of a proton from the hydroxyoxazolydine oxygen to the ring nitrogen) does not decompose spontaneously. Instead, it needs to overcome a barrier of 4.1 kcal/mol to reach the ester product. The ester form lies higher in energy than the reactants by 7.0 kcal/mol. The hydrolysis of the ester form to give the products (Gly + Ser) has to pass through two more transition states. The first one is associated with the -NH_2 catalyzed nucleophilic addition of H_2O and the second with the release of the Ser moiety. Both transition states are linked by a tetrahedral intermediate state. It should be noted that the calculated relative energies of these two TSs is lower by only 1.0 and 1.3 kcal/mol, respectively, as compared to the rate-determining TS.

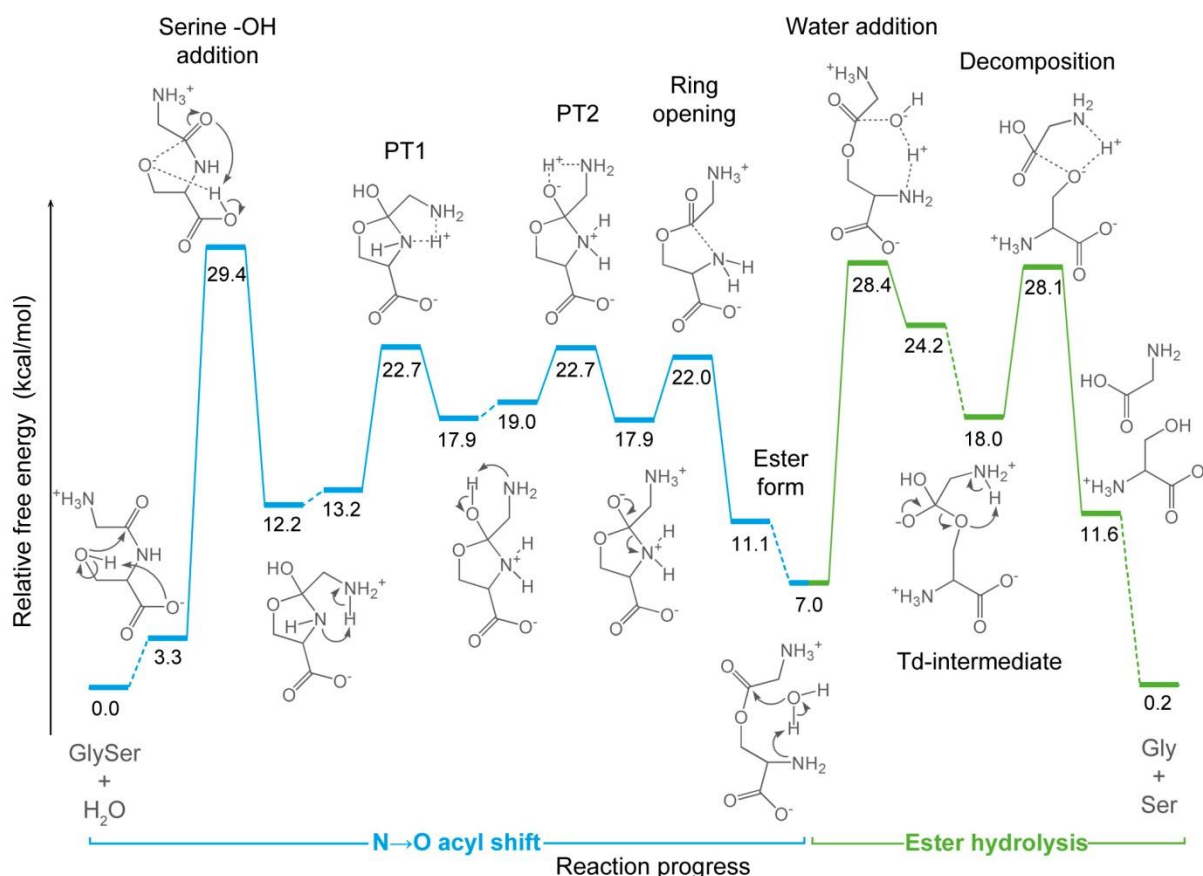


Figure 9. MP2 calculated free energy profile (kcal/mol) of the GlySer hydrolysis in neutral aqueous solution. The energy levels corresponding to structures successively obtained by following the intrinsic reaction coordinate are connected with a solid line. The energy levels of structures linked through conformational interconversion^[37] are connected with a dashed line. Some of the water molecules in the structural formulas representing the reaction process are omitted for simplicity. For more detailed structures we refer to Figure 8.

Protein self-cleavage at serine residues

As the calculations revealed, the lowest energy pathway for initiating the N→O acyl shift in GlySer involves the nucleophilic addition of the serine alcoholic function to the amide carbon, catalyzed by a proton transfer to the carboxylic group (TS-A1-1w, see Figure 3). Unless serine is located at the C-terminal position of the protein backbone this mechanism is not

applicable for peptides, where both the carboxylic and the amino groups are involved in a peptide link. One might assume, however, that properly positioned free carboxylic groups of aspartic or glutamic acid residues could promote the nucleophilic attack through a similar mechanism. It is interesting to compare the relative free energies of the transition states involved in channels A1, B1, and C1, where only the principle zwitterionic form is used (Figures. 3, 4 and 5). Let us recall that channel A, B and C exploit a PT from the serine -OH group to an O-atom of the carboxylic group, to the N-atom of the amide group and to the O-atom of the amide group, respectively. It is known that the amide oxygen is the preferred protonation site of amide groups. For instance, the energy difference between O- and N-protonated forms of N-methylacetamide was previously calculated in water (with the MP2 method and an IPCM solvation model) to be 7.2 kcal/mol in favor the former form.^[38] In case of GlySer we found that the O-protonated isomer is preferred over the N-protonated form by 11.2 kcal/mol, and less preferred with respect to the neutral form (protonated -COO⁻ group) by 7.5 kcal/mol (the corresponding lowest energy isomer conformations are given in Figure S6). Comparison of the energy requirements of channels A1, B1 and C1, shown in Figure 10(a), indicates that the TS energies decrease in the order channel B1 > channel C1 > channel A1, regardless of the number of solvent molecules included in the TS models (see also Figures S2–S4). This ordering follows the preference of protonation of different sites in GlySer: amide N-site < amide O-site < carboxyl O-site. Figure 10(b) reveals an almost linear relation between the activation energy of the most probable TSs of channels B1 (TS-B1-2w, $\Delta G_{aq}^{*,\ddagger} = 37.3$ kcal/mol), C1 (TS-C1-1w, $\Delta G_{aq}^{*,\ddagger} = 33.6$ kcal/mol) and A1 (TS-A1-1w, $\Delta G_{aq}^{*,\ddagger} = 29.4$ kcal/mol), and the relative free energy of the corresponding amide N-, amide O- and carboxyl O-protonated isomers of the principle zwitterion. Thus, it seems evident to conclude that the barrier height for intramolecular -OH addition is primarily determined by the aqueous-phase basicity of the proton accepting group, while being less sensitive to the specificity of the PT mechanism (direct or water mediated, number of solvent molecules involved in the bridge). If the proper position of the alcoholic function toward both the amide carbon (active conformation) and the general base are not hindered, this finding should also be valid for proteins. Not surprisingly, the relatively strong imidazole base of histidine residue has often been recognized as a general base catalyst for the N→X acyl shift in self-cleaving protein domains.^[11b, 39]

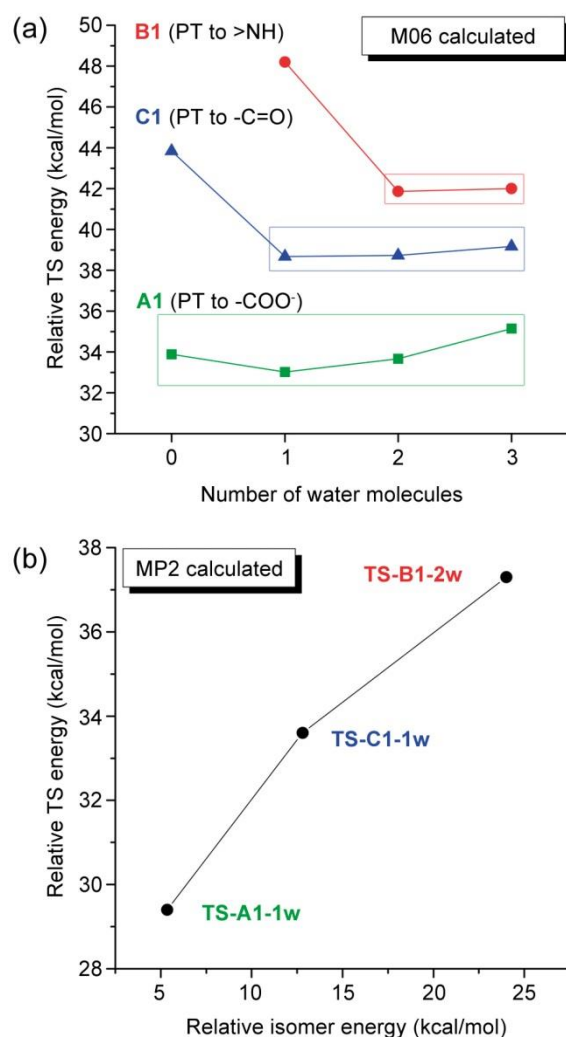


Figure 10. (a) Relative M06 calculated free energy of the TSs modeled within A1, B1 and C1 reaction channels as a function of the number of explicit water molecules. (b) Relative MP2 free energy of the most probable TSs of A1, B1 and C1 channels (TS-A1-1w, TS-B1-2w and TS-C1-1w, respectively) plotted as a function of the relative (with respect to the principle zwitterionic form) free energy of the corresponding carboxyl O-, amide N- and amide O-protonated isomers of the principle zwitterionic form. The molecular structures of the lowest energy isomers are given in Figure S6.

In contrast to A1, channels C1 and B1 do not require the assistance of an available proton accepting group belonging to an amino acid side chain in the vicinity of the scissile peptide bond. As estimated for the model compound, GlySer, the lower energy C1 channel passes through a barrier of 33.6 kcal/mol (TS-C1-1w; see Figure 7). Comparing this value with the activation energy of the spontaneous (uncatalyzed) hydrolysis of an internal peptide bond in the model compound acetylglycylglycine N-methylamide, ~32 kcal/mol (in H₂O at 25°C and pH 6.8)^[1], where the advantageous intramolecular -OH nucleophile is missing, it is obvious that in absence of additional catalytic factors the peptide bond adjacent to serine is no longer a weak point in the protein backbone. Hence, the presence of a general base seems to be crucial to unlock the nucleophilic ability of the Ser alcoholic group. This is one of the major points that distinguish the Ser -OH nucleophile from the Cys -SH one. As known for

protein splicing, in inteins utilizing a Cys residue as a nucleophile, the thiol group does not require assistance of other groups.^[40]

Alternatively, in the absence of a general base catalyst, the activation energy of a C1 like TS in proteins could be lowered through destabilization of the reactant state, as is the case of SEA domain autoproteolysis. The SEA domain is suggested to enhance the rate of proteolysis by straining the conformation of the amide plane of the scissile peptide bond, thus increasing the reactant free energy by ~7 kcal/mol.^[8] The activation free energy for the spontaneous proteolysis of GPR116 SEA and MUC1 SEA was experimentally estimated of ~21 kcal/mol.^[8a] Let us assume that in these proteins the N→O acyl shift is initiated through a TS-C1-1w (33.6 kcal/mol) mechanism which, in this case, operates on a strained peptide bond. After subtracting a ground state destabilization of ~7 kcal/mol (due to conformational stress) the required activation energy for intramolecular –OH attack in SEA becomes ~27 kcal/mol, which is ~6 kcal/mol above the experimental value. In a combined experimental and theoretical study Johansson et al. examined the pH dependence of SEA domain autoproteolysis.^[19a] The results suggested that the pK_a values for amide N- and amide O-protonation respond in opposite ways to strain, and nitrogen becomes the preferred site for protonation at a moderate nonplanarity (20°). Furthermore, when destroying the amide planarity (by applying constraints) such that it approaches the N-protonated conformation the pK_a shifts to the physiological range. Therefore, it was assumed that protonation, preceding the –OH nucleophilic attack, occurs at the amide nitrogen of the scissile peptide bond in SEA. Such a scenario closely resembles our C2 channel (TS-C2-1w) where the intramolecular nucleophilic addition, accompanied with a PT from the attacking –OH group to the amide oxygen, takes place in the amide N-protonated isomer of GlySer (Figure 5). By using the amide N-protonated isomer rather than the principle zwitterionic form as a reference for the C2 channel, the calculated activation energy of TS-C2-1w is reduced from 38.3 kcal/mol to 14.3 kcal/mol, which is 19.3 kcal/mol lower than for the unprotonated case, TS-C1-1w (33.6 kcal/mol). The predicted value of 14.3 kcal/mol is comparable to the experimental activation energy for acid-catalyzed hydrolysis of some distorted amides, 15 - 17 kcal/mol,^[41] but is much lower than the activation energy reported for the spontaneous proteolysis of SEA (~21 kcal/mol). A TS structure, similar to our TS-C2-1w was previously proposed by Johansson et al. (for a simplified model of GlySer, utilizing the B3LYP functional and CPCM solvation model). The calculated relative TS energy, however, was found to be rather high, ~40 kcal/mol. This could be explained by the fact that in the proposed TS a direct, rather than water mediated, PT from the Ser –OH to the amide oxygen was considered. As our calculations revealed, the activation energy of the C2 channel lowers dramatically (~17 kcal/mol according to the M06 calculations, Figure S4) when the PT is mediated by a water molecule.

There is still no clear evidence whether the intramolecular nucleophilic attack in SEA is preceded by protonation of the amide nitrogen or not. Apparently, more thorough theoretical investigations on this topic are needed. However, our results clearly show that the activation energy of the N→O acyl shift is significantly lowered when the H⁺ transfer from the attacking Ser-OH group to the amide oxygen is mediated by a water molecule (channels C1 and C2), regardless of the protonation state of the amide nitrogen. Hence, in both cases (protonated or not), taking into account that the cleavage site of SEA is entirely water-exposed,^[8b] the assistance of a bridging water molecule in the actual TS is strongly suggested. Admittedly, our calculations on this small model (GlySer) cannot by themselves provide all answers for the hydrolysis mechanism of the peptide bond adjacent to Ser in proteins, but we hope they provide some good directives.

Conclusions

To examine the inherent hydrolytic ability of serine-containing peptides, we have performed a detailed mechanistic study of the spontaneous hydrolysis of GlySer in neutral aqueous solution at the SMD-M06/6-311+G(2df,2p) level of theory. In order to identify the lowest energy path for the rate-determining step of GlySer hydrolysis four main reaction channels were considered, each of them representing one to three distinct pathways. In all cases the most important specific solute-solvent interactions were accounted for by explicit inclusion of water molecules in the models and by varying their number until a minimum in the relative TS energy was found. In the lowest energy transition state, TS -A1-1w, the serine COO⁻ group acts as a general base to accept a proton from the attacking -OH function, resulting in oxyoxazolidine ring closure. This process is assisted by one water molecule, hydrogen bonded to both the alcoholate oxygen and to an -NH₃⁺ hydrogen.

To check the reliability of the M06 functional in predicting the reaction energetics the relative energy of TS-A1-1w was also computed by means of the more accurate double-hybrid functionals (B2PLYP-D and mPW2PLYP-D) and with the MP2 method, and the results were compared to the experimental data obtained in this work. The barrier for the rate-determining step obtained from MP2 (29.4 kcal/mol) was found to be in closest agreement with the experimentally observed activation Gibbs function (29.4 kcal/mol). Therefore, this method was further used to compute the energy profile for the entire process of hydrolysis (making use of M06 structures). The proposed mechanism for GlySer hydrolysis involves a general-base catalyzed nucleophilic addition of the serine -OH site chain at the amide carbon, with the formation of a transient oxyoxazolidine ring. Decomposition of this state to an ester form is preceded by two proton transfers. The hydrolysis of the ester form involves an -NH₂ catalyzed water addition with formation of a tetrahedral intermediate followed by the release of the leaving serine moiety.

A comparative analysis of the activation energies of the TS models exploiting a PT from the attacking -OH group to the amide N-, amide O- and carboxyl O-atom in the principle zwitterion form (channels A1, B1 and C1) revealed that the $\Delta G_{aq}^{*,\#}$ values are primarily determined by the aqueous-phase basicity of the proton accepting group, while being less sensitive to the specificity of the PT mechanism. Since in proteins the Ser -COO⁻ group is involved in a peptide link A1-like TSs can no longer be obtained. In this case the second most probable candidate is channel C1 (involving a water-mediated PT from the Ser -OH to the amide oxygen, TS-C1-1w) which passes through a barrier of 33.6 kcal/mol. This barrier, however, is relatively high indicating that in absence of additional catalytic factors the peptide bond adjacent to serine is no longer a weak point in the protein backbone. In all our models involving a PT from the attacking nucleophile to the amide carbon or to the amide nitrogen, the activation energy for the N→O acyl shift lowers significantly when moving from direct to water-mediated PT. Thus, when considering peptide bond cleavage at the serine residue in absence of general base catalysts (assuming that the cleavage site is solvent-accessible) the participation of the water solvent in the transition state might be crucial for the observed reaction rate.

Acknowledgments

We acknowledge the Flemish Science Foundation (FWO) for the financial support under project G.0260.12. All calculations were performed on the HPC cluster VIC of the KU Leuven.

Supporting information

Additional figures, S1 – S6. Optimized molecular geometries in Cartesian coordinates.

References

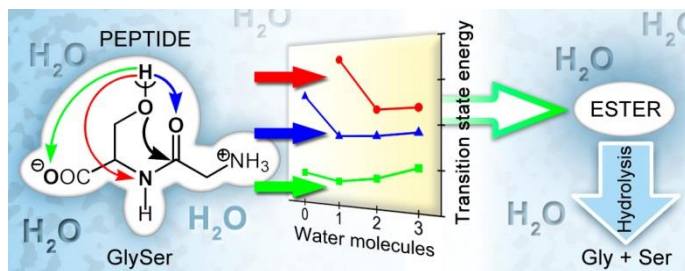
- [1] A. Radzicka, R. Wolfender, *J. Am. Chem. Soc.* **1996**, *118*, 6105-6109.
- [2] F. B. Perler, M.-Q. Xu, H. Paulus, *Curr. Opin. Chem. Biol.* **1997**, *1*, 292-299.
- [3] C. J. Noren, J. Wang, F. B. Perler, *Angew. Chem. Int. Ed.* **2000**, *39*, 450-466.
- [4] P. D. van Poelje, E. E. Snell, *Annu. Rev. Biochem.* **1990**, *59*, 29-59.
- [5] a) J. A. Porter, D. P. v. Kessler, S. C. Ekker, K. E. Young, J. J. Lee, K. Moses, P. A. Beachy, *Nature* **1996**, *374*, 363 - 366; b) J. A. Porter, K. E. Young, P. A. Beachy, *Science* **1996**, *274*, 255-259.
- [6] J. A. Brannigan, G. Dodson, H. J. Duggleby, P. C. E. Moody, J. L. Smith, D. R. Tomchick, A. G. Murzin, *Nature* **1995**, *378*, 416 - 419.
- [7] A. E. Hodel, M. R. Hodel, E. R. Griffis, K. A. Hennig, G. A. Ratner, S. Xu, M. A. Powers, *Mol. Cell* **2002**, *10*, 347-358.
- [8] a) A. Sandberg, D. G. A. Johansson, B. Macao, *J. Mol. Biol.* **2008**, *377*, 1117-1129; b) D. G. A. Johansson, B. Macao, A. Sandberg, T. Härd, *J. Mol. Biol.* **2008**, *377*, 1130-1143.
- [9] H. Paulus, *Chem. Soc. Rev.* **1998**, *27*, 375-386.
- [10] K. Iwai, T. Ando, *Methods Enzymol.* **1967**, *11*, 263-282.
- [11] a) K. Michalska, A. Hernandez-Santoyo, M. Jaskolski, *J. Biol. Chem.* **2008**, *283*, 13388-13397; b) Z. Du, P. T. Shemella, Y. Liu, S. A. McCallum, B. Pereira, S. K. Nayak, G. Belfort, M. Belfort, C. Wang, *J. Am. Chem. Soc.* **2009**, *131*, 11581-11589; c) Z. Du, Y. Zheng, M. Patterson, Y. Liu, C. Wang, *J. Am. Chem. Soc.* **2011**, *133*, 10275-10282.
- [12] Q. Xu, D. Buckley, C. Guan, H.-C. Guo, *Cell* **1999**, *98*, 651-661.
- [13] a) T. Klabunde, S. Sharma, A. Telenti, W. R. Jacobs, Jr., J. C. Sacchettini, *Nat. Struct. Biol.* **1998**, *5*, 31-36; b) B. W. Poland, M.-Q. Xu, F. A. Quiocho, *J. Biol. Chem.* **2000**, *275*, 16408-16413.
- [14] C. A. Lewis, Jr., R. Wolfenden, *Biochemistry* **2011**, *50*, 7259-7264.
- [15] a) P. H. Ho, K. Stroobants, T. N. Parac-Vogt, *Inorg. Chem.* **2011**, *50*, 12025-12033; b) P. H. Ho, T. Mihaylov, K. Pierloot, T. N. Parac-Vogt, *Inorg. Chem.* **2012**, *51*, 8848.
- [16] J. A. Dean, 15th ed., McGraw-Hill, Inc., New York, **1999**.
- [17] E. Bottari, D. Cellulosi, M. R. Festa, *Talanta* **1999**, *50*, 993-1002.
- [18] M. Yashiro, Y. Sonobe, A. Yamamura, T. Takarada, M. Komiyama, Y. Fujii, *Org. Biomol. Chem.* **2003**, *1*, 629-632.
- [19] a) D. G. A. Johansson, G. Wallin, A. Sandberg, B. Macao, J. Åqvist, T. Härd, *J. Am. Chem. Soc.* **2009**, *131*, 9475-9477; b) G. Wallin, T. Härd, J. Åqvist, *J. Chem. Theory Comput.* **2012**, *8*, 3871-3879.

- [20] R. B. Sunoj, M. Anand, *Phys. Chem. Chem. Phys.* **2012**, *14*, 12715–12736.
- [21] L. Bethencourt, O. Nunez, *J. Org. Chem.* **2008**, *73*, 2105–2113.
- [22] J. R. Pliego, Jr., J. M. Riveros, *Chem. Eur. J.* **2002**, *8*, 1945–1953.
- [23] Y. Zhao, D. G. Truhlar, *Theor. Chem. Account* **2008**, *120*, 215–241.
- [24] A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378–6396.
- [25] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- [26] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.
- [27] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108–034116.
- [28] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- [29] a) J. I. Mujika, J. M. Mercero, X. Lopez, *J. Am. Chem. Soc.* **2005**, *127*, 4445–4453; b) J. Pitarch, M. F. Ruiz-López, E. Silla, J.-L. Pascual-Ahuir, I. Tuñón, *J. Am. Chem. Soc.* **1998**, *120*, 2146–2155; c) L. Gorb, A. Asensio, I. Tuñón, M. F. Ruiz-López, *Chem. Eur. J.* **2005**, *11*, 6743–6753; d) G. I. Almerindo, J. R. Pliego, *J. Braz. Chem. Soc.* **2007**, *18*, 696–702; e) A. Kumar, X. Zhu, K. Walsh, R. Prabhakar, *Inorg. Chem.* **2010**, *49*, 38–46; f) R. P. Bora, A. Barman, X. Zhu, M. Ozbil, R. Prabhakar, *J. Phys. Chem. B* **2010**, *114*, 10860–10875; g) M. Topf, W. G. Richards, *J. Am. Chem. Soc.* **2004**, *126*, 14631–14641.
- [30] a) B. Wang, Z. Cao, *J. Phys. Chem. A* **2010**, *114*, 12918–12927; b) J. A. Tossell, *Inorg. Chem.* **2006**, *45*, 5961–5970; c) Y. K. Kang, *J. Phys. Chem. B* **2007**, *111*, 10550–10556; d) X. Chen, C. K. Regan, S. L. Craig, E. H. Krenske, K. N. Houk, W. L. Jorgensen, J. I. Brauman, *J. Am. Chem. Soc.* **2009**, *131*, 16162–16170; e) J. L. Przybylski, S. D. Wetmore, *J. Phys. Chem. B* **2009**, *113*, 6533–6542; f) J. L. Przybylski, S. D. Wetmore, *J. Phys. Chem. B* **2010**, *114*, 1104–1113; g) B. Wang, Z. Cao, *Chem. Eur. J.* **2011**, *17*, 11919–11929.
- [31] J. Ho, A. Klamt, M. L. Coote, *J. Phys. Chem. A* **2010**, *114*, 13442–13444.
- [32] R. F. Ribeiro, A. V. Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2011**, *115*, 14556–14562.
- [33] J. R. Pliego, Jr., J. M. Riveros, *J. Phys. Chem. A* **2001**, *105*, 7241–7247.
- [34] V. S. Bryantsev, M. S. Diallo, W. A. Goddard III, *J. Phys. Chem. B* **2008**, *112*, 9709–9719.
- [35] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Revision A.02 ed., Gaussian, Inc., Wallingford CT, **2009**.
- [36] B. R. Martin, *Biopolymers* **1998**, *45*, 351–353.
- [37] In order to find the most probable reaction mechanism our efforts were mainly focused at finding the lowest energy TSs, involving bond breaking and/or bond formation, which might occur during the hydrolysis reaction. Every TS found bridges two local minima (intermediate structures) obtained by following the intrinsic reaction coordinate in the forward and in the reverse direction respectively. Thus, the two local minima between two consecutive TSs (each of which is derived either from one or

from the other TS) are usually in slightly different conformations and to be linked in the reaction path the conformational interconversion should in principle also be tracked. However, as they involve relatively unhindered single bond rotations we may safely assume that the energy barriers of conformational interconversion are smaller than the barriers associated with bond breaking/formation and will not affect the general picture of the reaction energy profile.

- [38] A. Bagno, B. Bujnicki, S. Bertrand, C. Comuzzi, F. Dorigo, P. Janvier, G. Scorrano, *Chem. Eur. J.* **1999**, 5, 523-536.
- [39] L. Saleh, F. B. Perler, *Chem. Record* **2006**, 6, 183–193.
- [40] C. J. Noren, J. Wang, F. B. Perler, *Angew. Chem. Int. Ed. Engl.* **2000**, 39, 450-466.
- [41] Q.-P. Wang, A. J. Bennet, R. S. Brown, B. D. Santarsiero, *J. Am. Chem. Soc.* **1991**, 113, 5757-5765.

Graphical abstract



Protein self-cleavage at serine residues: To examine the inherent hydrolytic ability of serine-containing peptides we have performed a detailed theoretical and experimental study of the spontaneous hydrolysis of glycylserine. Based on our models, possible mechanistic scenarios for both uncatalyzed and self-catalyzed peptide bond cleavage in proteins have been proposed.